MODERATE DEVIATIONS FOR THE DURBIN-WATSON STATISTIC RELATED TO THE FIRST-ORDER AUTOREGRESSIVE PROCESS

S.VALÈRE BITSEKI PENDA, HACÈNE DJELLOUT, AND FRÉDÉRIC PROÏA

Abstract. The purpose of this paper is to investigate moderate deviations for the Durbin-Watson statistic associated with the stable first-order autoregressive process where the driven noise is also given by a first-order autoregressive process. We first establish a moderate deviation principle for both the least squares estimator of the unknown parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. It enables us to provide a moderate deviation principle for the Durbin-Watson statistic in the easy case where the driven noise is normally distributed and in the more general case where the driven noise satisfies a less restrictive Chen-Ledoux type condition.

AMS 2000 subject classifications: 60F10, 60G42, 62M10, 62G05.

1. Introduction

This paper is focused on the stable first-order autoregressive process where the driven noise is also given by a first-order autoregressive process. The purpose is to investigate moderate deviations for both the least squares estimator of the unknown parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. Our goal is to establish moderate deviations for the Durbin-Watson statistic [10], [11], [12], in a lagged dependent random variables framework. First of all, we shall assume that the driven noise is normally distributed. Then, we will extend our investigation to the more general framework where the driven noise satisfies a less restrictive Chen-Ledoux type condition [4], [16]. We are inspired by the recent paper of Bercu and Proïa [2], where the almost sure convergence and the central limit theorem are established for both the least squares estimators and the Durbin-Watson statistic. Our results are proved via an extensive use of the results of Dembo [5], Dembo and Zeitouni [6] and Worms [22], [23], [24] on the one hand, and of the paper of Puhalskii [19] and Djellout [7] on the other hand, about moderate deviations for martingales. In order to introduce the Durbin-Watson statistic, we shall focus our attention on the first-order autoregressive process given, for all $n \geq 1$, by

$$\left\{ \begin{array}{rcll} X_n & = & \theta X_{n-1} & + \quad \varepsilon_n \\ \varepsilon_n & = & \rho \varepsilon_{n-1} & + \quad V_n \end{array} \right. \tag{1.1}$$

where we shall assume that the unknown parameters $|\theta| < 1$ and $|\rho| < 1$ to ensure the stability of the model. In all the sequel, we also assume that (V_n) is a sequence of independent and identically distributed random variables with zero mean, positive variance σ^2 and satisfying some suitable assumptions. The square-integrable initial values X_0 and ε_0 may

Key words and phrases. Durbin-Watson statistic, Moderate deviation principle, First-order autoregressive process, Serial correlation.

be arbitrarily chosen. We have decided to estimate θ by the least squares estimator

$$\widehat{\theta}_n = \frac{\sum_{k=1}^{n} X_k X_{k-1}}{\sum_{k=1}^{n} X_{k-1}^2}. \tag{1.2}$$

Then, we also define a set of least squares residuals given, for all $1 \leq k \leq n$, by

$$\widehat{\varepsilon}_k = X_k - \widehat{\theta}_n X_{k-1}, \tag{1.3}$$

which leads to the estimator of ρ,

$$\widehat{\rho}_n = \frac{\sum_{k=1}^{n} \widehat{\varepsilon}_k \widehat{\varepsilon}_{k-1}}{\sum_{k=1}^{n} \widehat{\varepsilon}_{k-1}^{\,2}}. \tag{1.4}$$

Finally, the Durbin-Watson statistic is defined, for $n \geq 1$, as

$$\widehat{D}_n = \frac{\sum_{k=1}^{n} (\widehat{\varepsilon}_k - \widehat{\varepsilon}_{k-1})^2}{\sum_{k=0}^{n} \widehat{\varepsilon}_{k}^{\,2}}. \tag{1.5}$$

This well-known statistic was introduced by the pioneer work of Durbin and Watson [10], [11], [12], in the middle of last century, to test the presence of a significative first order serial correlation in the residuals of a regression analysis. A wide range of litterature is available on the asymptotic behavior of the Durbin-Watson statistic, frequently used in Econometry. While it appeared to work pretty well in the classical independent framework, Malinvaud [17] and Nerlove and Wallis [18] observed that, for linear regression models containing lagged dependent random variables, the Durbin-Watson statistic may be asymptotically biased, potentially leading to inadequate conclusions. Durbin [9] proposed alternative tests to prevent this misuse, such as the *h-test* and the *t-test*, then substantial contributions were brought by Inder [14], King and Wu [15] and more recently Stocker [20]. Lately, a set of results have been established by Bercu and Proïa in [2], in particular a test procedure as powerful as the *h-test*, and they will be summarized thereafter as a basis for this paper.

The paper is organized as follows. First of all, we recall the results recently established by Bercu and Proïa [2]. In Section 2, we propose moderate deviation principles for the estimators of θ and ρ and for the Durbin-Watson statistic, given by (1.2), (1.4) and (1.5), under the normality assumption on the driven noise. Section 3 deals with the generalization of the latter results under a less restrictive Chen-Ledoux type condition on (V_n). Finally, all technical proofs are postponed to Section 4.

Lemma 1.1. *We have the almost sure convergence of the autoregressive estimator,*

$$\lim_{n\to\infty} \widehat{\theta}_n = \theta^* \quad \text{a.s.}$$

where the limiting value

$$\theta^* = \frac{\theta + \rho}{1 + \theta\rho}. \tag{1.6}$$

In addition, as soon as $\mathbb{E}[V_1^4] < \infty$, we also have the asymptotic normality,

$$\sqrt{n}\left(\widehat{\theta}_n - \theta^*\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\theta^2)$$

where the asymptotic variance

$$\sigma_\theta^2 = \frac{(1-\theta^2)(1-\theta\rho)(1-\rho^2)}{(1+\theta\rho)^3}. \tag{1.7}$$

Lemma 1.2. *We have the almost sure convergence of the serial correlation estimator,*

$$\lim_{n\to\infty} \widehat{\rho}_n = \rho^* \quad \text{a.s.}$$

where the limiting value

$$\rho^* = \theta\rho\theta^*. \tag{1.8}$$

Moreover, as soon as $\mathbb{E}[V_1^4] < \infty$*, we have the asymptotic normality,*

$$\sqrt{n}\Big(\widehat{\rho}_n - \rho^*\Big) \xrightarrow{\mathcal{L}} \mathcal{N}\big(0, \sigma_\rho^2\big)$$

with the asymptotic variance

$$\sigma_\rho^2 = \frac{(1-\theta\rho)}{(1+\theta\rho)^3}\big((\theta+\rho)^2(1+\theta\rho)^2 + (\theta\rho)^2(1-\theta^2)(1-\rho^2)\big). \tag{1.9}$$

On top of that, we have the joint asymptotic normality,

$$\sqrt{n}\begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\big(0, \Gamma\big)$$

where the covariance matrix

$$\Gamma = \begin{pmatrix} \sigma_\theta^2 & \theta\rho\sigma_\theta^2 \\ \theta\rho\sigma_\theta^2 & \sigma_\rho^2 \end{pmatrix}. \tag{1.10}$$

Lemma 1.3. *We have the almost sure convergence of the Durbin-Watson statistic,*

$$\lim_{n\to\infty} \widehat{D}_n = D^* \quad \text{a.s.}$$

where the limiting value

$$D^* = 2(1-\rho^*). \tag{1.11}$$

In addition, as soon as $\mathbb{E}[V_1^4] < \infty$*, we have the asymptotic normality,*

$$\sqrt{n}\left(\widehat{D}_n - D^*\right) \xrightarrow{\mathcal{L}} \mathcal{N}\big(0, \sigma_D^2\big)$$

where the asymptotic variance

$$\sigma_D^2 = 4\sigma_\rho^2. \tag{1.12}$$

Proof. The proofs of Lemma 1.1, Lemma 1.2 and Lemma 1.3 may be found in [2]. □

Our objective is to establish a set of moderate deviation principles on these estimates in order to get a better asymptotic precision than the central limit theorem. In all the sequel, (b_n) will denote a sequence of increasing positive numbers satisfying $1 = o(b_n^2)$ and $b_n^2 = o(n)$, that is

$$b_n \longrightarrow \infty, \qquad \frac{b_n}{\sqrt{n}} \longrightarrow 0. \tag{1.13}$$

Remarks and Notations. *In the whole paper, for any matrix* M*,* M' *and* $\|M\|$ *stand for the transpose and the euclidean norm of* M*, respectively. For any square matrix* M*,* $\det(M)$ *and* $\rho(M)$ *are the determinant and the spectral radius of* M*, respectively. Moreover, we will shorten* large deviation principle *by LDP. In addition, for a sequence of random variables*

$(Z_n)_n$ *on* $\mathbb{R}^{d\times p}$*, we say that* $(Z_n)_n$ *converges* $(b_n^2)-$*superexponentially fast in probability to some random variable* Z *if, for all* $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\|Z_n - Z\| > \delta \Big) = -\infty.$$

This exponential convergence *with speed* b_n^2 *will be shortened as*

$$Z_n \overset{\text{superexp}}{\underset{b_n^2}{\longrightarrow}} Z.$$

The exponential equivalence *with speed* b_n^2 *between two sequences of random variables* $(Y_n)_n$ *and* $(Z_n)_n$*, whose precise definition is given in Definition 4.2.10 of* [6]*, will be shortened as*

$$Y_n \overset{\text{superexp}}{\underset{b_n^2}{\sim}} Z_n.$$

2. On moderate deviations under the Gaussian condition

In this first part, we focus our attention on moderate deviations for the Durbin-Watson statistic in the easy case where the driven noise (V_n) is normally distributed. This restrictive assumption allows us to reduce the set of hypothesis to the existence of $t > 0$ such that

(G.1)

$$\mathbb{E}\Big[\exp(t\varepsilon_0^2) \Big] < \infty,$$

(G.2)

$$\mathbb{E}\Big[\exp(tX_0^2) \Big] < \infty.$$

Theorem 2.1. *Assume that there exists* $t > 0$ *such that* **(G.1)** *and* **(G.2)** *are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n} \left(\widehat{\theta}_n - \theta^* \right) \right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}$ *with speed* b_n^2 *and good rate function*

$$I_\theta(x) = \frac{x^2}{2\sigma_\theta^2} \tag{2.1}$$

where σ_θ^2 *is given by* (1.7).

Theorem 2.2. *Assume that there exists* $t > 0$ *such that* **(G.1)** *and* **(G.2)** *are satisfied. Then, as soon as* $\theta \neq -\rho$*, the sequence*

$$\left(\frac{\sqrt{n}}{b_n} \begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\rho}_n - \rho^* \end{pmatrix} \right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}^2$ *with speed* b_n^2 *and good rate function*

$$K(x) = \frac{1}{2} x' \Gamma^{-1} x \tag{2.2}$$

where Γ *is given by* (1.10)*. In particular, the sequence*

$$\left(\frac{\sqrt{n}}{b_n} \left(\widehat{\rho}_n - \rho^* \right) \right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}$ with speed b_n^2 and good rate function

$$I_\rho(x) = \frac{x^2}{2\sigma_\rho^2} \tag{2.3}$$

where σ_ρ^2 is given by (1.9).

Remark 2.1. *The covariance matrix Γ is invertible if and only if $\theta \neq -\rho$ since one can see by a straightforward calculation that*

$$\det(\Gamma) = \frac{\sigma_\theta^2(\theta+\rho)^2(1-\theta\rho)}{(1+\rho^2)}.$$

Moreover, in the particular case where $\theta = -\rho$, the sequences

$$\left(\frac{\sqrt{n}}{b_n}\left(\widehat{\theta}_n - \theta^*\right)\right)_{n\geq 1} \qquad \text{and} \qquad \left(\frac{\sqrt{n}}{b_n}\left(\widehat{\rho}_n - \rho^*\right)\right)_{n\geq 1}$$

satisfy LDP on $\mathbb{R}$ with speed b_n^2 and good rate functions respectively given by

$$I_\theta(x) = \frac{x^2(1-\theta^2)}{2(1+\theta^2)} \qquad \text{and} \qquad I_\rho(x) = \frac{x^2(1-\theta^2)}{2\theta^4(1+\theta^2)}.$$

Theorem 2.3. *Assume that there exists $t > 0$ such that* **(G.1)** *and* **(G.2)** *are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}\left(\widehat{D}_n - D^*\right)\right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}$ with speed b_n^2 and good rate function

$$I_D(x) = \frac{x^2}{2\sigma_D^2} \tag{2.4}$$

where σ_D^2 is given by (1.12).

Proof. Theorem 2.1, Theorem 2.2 and Theorem 2.3 are proved in Section 4. □

3. On moderate deviations under the Chen-Ledoux type condition

Via an extensive use of Puhalskii's result, we will now focus our attention on the more general framework where the driven noise (V_n) is assumed to satisfy the Chen-Ledoux type condition. Accordingly, one shall introduce the following hypothesis, for $a = 2$ and $a = 4$.

(CL.1) Chen-Ledoux.

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log n\mathbb{P}\Big(|V_1|^a > b_n\sqrt{n}\Big) = -\infty.$$

(CL.2)

$$\frac{|\varepsilon_0|^a}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0.$$

(CL.3)

$$\frac{|X_0|^a}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0.$$

Remark 3.1. *If the random variable V_1 satisfies* **(CL.1)** *with $a = 2$, then*

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log n\mathbb{P}\Big(\big|V_1^2 - \mathbb{E}[V_1^2]\big| > b_n\sqrt{n}\Big) = -\infty, \tag{3.1}$$

which implies in particular that $\mathrm{Var}(V_1^4) < \infty$. *Moreover, if the random variable V_1 has exponential moments, i.e. if there exists $t > 0$ such that*

$$\mathbb{E}\Big[\exp\left(tV_1^2\right)\Big] < \infty,$$

then **(CL.1)** *is satisfied for every increasing sequence (b_n). From* [1], [13], *condition* (3.1) *is equivalent to say that the sequence*

$$\left(\frac{1}{b_n\sqrt{n}} \sum_{k=1}^{n} \Big(V_k^2 - \mathbb{E}[V_k^2]\Big)\right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}$ with speed b_n^2 and good rate function

$$I(x) = \frac{x^2}{2\mathrm{Var}(V_1^2)}.$$

Remark 3.2. *If we choose $b_n = n^\alpha$ with $0 < \alpha < 1/2$,* **(CL.1)** *is immediately satisfied if there exists $t > 0$ and $0 < \beta < 1$ such that*

$$\mathbb{E}\Big[\exp\left(tV_1^{2\beta}\right)\Big] < \infty,$$

which is clearly a weaker assumption than the existence of $t > 0$ such that

$$\mathbb{E}\Big[\exp\left(tV_1^2\right)\Big] < \infty,$$

imposed in the previous section.

Remark 3.3. *If* **(CL.1)** *is satisfied for $a = 4$, then it is also satisfied for all $0 < b < a$.*

Remark 3.4. *In the technical proofs that will follow, rather than* **(CL.1)** *with $a = 4$, the weakest assumption really needed could be summarized by the existence of a large constant C such that*

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} V_k^4 > C\right) = -\infty.$$

Theorem 3.1. *Assume that* **(CL.1)**, **(CL.2)** *and* **(CL.3)** *are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}\left(\widehat{\theta}_n - \theta^*\right)\right)_{n\geq 1}$$

satisfies the LDP on $\mathbb{R}$ given in Theorem 2.1.

Theorem 3.2. *Assume that* **(CL.1)**, **(CL.2)** *and* **(CL.3)** *are satisfied. Then, as soon as $\theta \neq -\rho$, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}\begin{pmatrix}\widehat{\theta}_n - \theta^* \\ \widehat{\rho}_n - \rho^*\end{pmatrix}\right)_{n\geq 1}$$

satisfies the LDP on $\mathbb{R}^2$ given in Theorem 2.2. In particular, the sequence

$$\left(\frac{\sqrt{n}}{b_n}\left(\widehat{\rho}_n - \rho^*\right)\right)_{n\geq 1}$$

satisfies the LDP on $\mathbb{R}$ also given in Theorem 2.2.

Remark 3.5. *We have already seen in Remark 2.1 that the covariance matrix Γ is invertible if and only if $\theta \neq -\rho$. In the particular case where $\theta = -\rho$, the sequences*

$$\left(\frac{\sqrt{n}}{b_n}\left(\widehat{\theta}_n - \theta^*\right)\right)_{n\geq 1} \qquad \text{and} \qquad \left(\frac{\sqrt{n}}{b_n}\left(\widehat{\rho}_n - \rho^*\right)\right)_{n\geq 1}$$

satisfy the LDP on $\mathbb{R}$ given in Remark 2.1.

Theorem 3.3. *Assume that* **(CL.1)**, **(CL.2)** *and* **(CL.3)** *are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}\left(\widehat{D}_n - D^*\right)\right)_{n\geq 1}$$

satisfies the LDP on $\mathbb{R}$ given in Theorem 2.3.

Proof. Theorem 3.1, Theorem 3.2 and Theorem 3.3 are proved in Section 4. □

4. Proof of the main results

For a matter of readability, some notations commonly used in the following proofs have to be introduced. First, for all $n \geq 1$, let

$$L_n = \sum_{k=1}^{n} V_k^2. \tag{4.1}$$

Then, let us define M_n, for all $n \geq 1$, as

$$M_n = \sum_{k=1}^{n} X_{k-1} V_k \tag{4.2}$$

where $M_0 = 0$. For all $n \geq 1$, denote by $\mathcal{F}_n$ the σ-algebra of the events occurring up to time n, $\mathcal{F}_n = \sigma(X_0, \varepsilon_0, V_1, \cdots, V_n)$. We infer from (4.2) that $(M_n)_{n\geq 0}$ is a locally square-integrable real martingale with respect to the filtration $\mathbb{F} = (\mathcal{F}_n)_{n\geq 0}$ with predictable quadratic variation given by $\langle M\rangle_0 = 0$ and for all $n \geq 1$, $\langle M\rangle_n = \sigma^2 S_{n-1}$, where

$$S_n = \sum_{k=0}^{n} X_k^2. \tag{4.3}$$

Moreover, $(N_n)_{n\geq 0}$ is defined, for all $n \geq 2$, as

$$N_n = \sum_{k=2}^{n} X_{k-2} V_k \tag{4.4}$$

and $N_0 = N_1 = 0$. It is not hard to see that $(N_n)_{n\geq 0}$ is also a locally square-integrable real martingale sharing the same properties than $(M_n)_{n\geq 0}$. More precisely, its predictable quadratic variation is given by $\langle N\rangle_n = \sigma^2 S_{n-2}$. To conclude, let $P_0 = 0$ and, for all $n \geq 1$,

$$P_n = \sum_{k=1}^{n} X_{k-1} X_k. \tag{4.5}$$

4.1. Proof of Theorem 2.1.

Before starting the proof of Theorem 2.1, we need to introduce some technical tools. Denote by ℓ the almost sure limit of S_n/n [2], given by

$$\ell = \frac{\sigma^2(1+\theta\rho)}{(1-\theta^2)(1-\theta\rho)(1-\rho^2)}. \tag{4.6}$$

Lemma 4.1. *Under the assumptions of Theorem 2.1, we have the exponential convergence*

$$\frac{S_n}{n} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} \ell \tag{4.7}$$

where ℓ is given by (4.6).

Proof. After straightforward calculations, we get that for all $n \geq 2$,

$$\frac{S_n}{n} - \ell = \frac{\ell}{\sigma^2}\left[\left(\frac{L_n}{n} - \sigma^2\right) + 2\theta^* \frac{M_n}{n} - 2\theta\rho\frac{N_n}{n} + \frac{R_n}{n}\right] \tag{4.8}$$

where L_n, M_n, S_n and N_n are respectively given by (4.1), (4.2), (4.3) and (4.4),

$$R_n = [2(\theta+\rho)\rho^* - (\theta+\rho)^2 - (\theta\rho)^2]X_n^2 - (\theta\rho)^2 X_{n-1}^2 + 2\rho^* X_n X_{n-1} + \xi_1,$$

and where the remainder term

$$\xi_1 = (1 - 2\theta\rho - \rho^2)X_0^2 + \rho^2\varepsilon_0^2 + 2\theta\rho X_0\varepsilon_0 - 2\rho\rho^*(\varepsilon_0 - X_0)X_0 + 2\rho(\varepsilon_0 - X_0)V_1.$$

First of all, (V_n) is a sequence of independent and identically distributed gaussian random variables with zero mean and variance $\sigma^2 > 0$. It immediately follows from Cramér-Chernoff's Theorem, expounded e.g. in [6], that for all $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left|\frac{L_n}{n} - \sigma^2\right| > \delta\right) < 0. \tag{4.9}$$

Since $b_n^2 = o(n)$, the latter convergence leads to

$$\frac{L_n}{n} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} \sigma^2, \tag{4.10}$$

ensuring the exponential convergence of L_n/n to σ^2 with speed b_n^2. Moreover, for all $\delta > 0$ and a suitable $t > 0$, we clearly obtain from Markov's inequality that

$$\mathbb{P}\left(\frac{X_0^2}{n} > \delta\right) \leq \exp\left(-tn\delta\right)\mathbb{E}\Big[\exp(tX_0^2)\Big],$$

which immediately implies via **(G.2)**,

$$\frac{X_0^2}{n} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} 0, \tag{4.11}$$

and we get the exponential convergence of X_0^2/n to 0 with speed b_n^2. The same is true for V_1^2/n, ε_0^2/n and more generally for any isolated term of order 2 in relation (4.8) whose numerator do not depend on n. Let us now focus our attention on X_n^2/n. The model (1.1) can be rewritten in the vectorial form,

$$\Phi_n = A\Phi_{n-1} + W_n \tag{4.12}$$

where $\Phi_n = \begin{pmatrix} X_n & X_{n-1} \end{pmatrix}'$ stands for the lag vector of order 2, $W_n = \begin{pmatrix} V_n & 0 \end{pmatrix}'$ and

$$A = \begin{pmatrix} \theta+\rho & -\theta\rho \\ 1 & 0 \end{pmatrix}. \tag{4.13}$$

It is easy to show that $\rho(A) = \max(|\theta|, |\rho|) < 1$ under the stability conditions. According to Proposition 4.1 of [22],

$$\frac{\|\Phi_n\|^2}{n} \underset{n}{\overset{\text{superexp}}{\longrightarrow}} 0,$$

which is clearly sufficient to deduce that

$$\frac{X_n^2}{n} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} 0. \tag{4.14}$$

The exponential convergence of R_n/n to 0 with speed b_n^2 is achieved following exactly the same lines. To conclude the proof of Lemma 4.1, it remains to study the exponential asymptotic behavior of M_n/n. For all $\delta > 0$ and a suitable $y > 0$,

$$\begin{aligned}
\mathbb{P}\left(\frac{M_n}{n} > \delta\right) &= \mathbb{P}\left(\frac{M_n}{n} > \delta, \langle M\rangle_n \leq y\right) + \mathbb{P}\left(\frac{M_n}{n} > \delta, \langle M\rangle_n > y\right), \\
&\leq \exp\left(-\frac{n^2\delta^2}{2y}\right) + \mathbb{P}\Big(\langle M\rangle_n > y\Big),
\end{aligned} \tag{4.15}$$

by application of Theorem 4.1 of [3] in the case of a gaussian martingale. Then, noting that we have the following inequality,

$$S_n \leq \alpha X_0^2 + \beta\varepsilon_0^2 + \beta L_n \quad \text{a.s.} \tag{4.16}$$

with $\alpha = 1 + (1-|\theta|)^{-2}$ and $\beta = (1-|\rho|)^{-2}(1-|\theta|)^{-2}$, we get for a suitable $t > 0$,

$$\begin{aligned}
\mathbb{P}\Big(\langle M\rangle_n > y\Big) &\leq \mathbb{P}\left(X_0^2 > \frac{y}{3\alpha\sigma^2}\right) + \mathbb{P}\left(\varepsilon_0^2 > \frac{y}{3\beta\sigma^2}\right) + \mathbb{P}\left(L_{n-1} > \frac{y}{3\beta\sigma^2}\right), \\
&\leq \exp\left(\frac{-yt}{3\alpha\sigma^2}\right)\mathbb{E}\Big[\exp(tX_0^2)\Big] + \exp\left(\frac{-yt}{3\beta\sigma^2}\right)\mathbb{E}\Big[\exp(t\varepsilon_0^2)\Big] \\
&\qquad + \mathbb{P}\left(L_{n-1} > \frac{y}{3\beta\sigma^2}\right), \\
&\leq 3\max\Bigg(\exp\left(\frac{-yt}{3\alpha\sigma^2}\right)\mathbb{E}\Big[\exp(tX_0^2)\Big], \exp\left(\frac{-yt}{3\beta\sigma^2}\right)\mathbb{E}\Big[\exp(t\varepsilon_0^2)\Big], \\
&\qquad \mathbb{P}\left(L_{n-1} > \frac{y}{3\beta\sigma^2}\right)\Bigg).
\end{aligned}$$

Let us choose $y = nx$, assuming $x > 3\beta\sigma^4$. It follows that

$$\begin{aligned}
\frac{1}{b_n^2}\log\mathbb{P}\Big(\langle M\rangle_n > nx\Big) \leq\ & \frac{\log 3}{b_n^2} + \frac{1}{b_n^2}\max\Bigg(\frac{-nxt}{3\alpha\sigma^2} + \log\mathbb{E}\Big[\exp(tX_0^2)\Big], \\
& \frac{-nxt}{3\beta\sigma^2} + \log\mathbb{E}\Big[\exp(t\varepsilon_0^2)\Big], \log\mathbb{P}\left(L_{n-1} > \frac{nx}{3\beta\sigma^2}\right)\Bigg).
\end{aligned}$$

Since $b_n^2 = o(n)$ and by virtue of (4.10) with $\delta = x/(3\beta\sigma^2) - \sigma^2 > 0$, we obtain that

$$\limsup_{n\to\infty} \frac{1}{b_n^2}\log\mathbb{P}\Big(\langle M\rangle_n > nx\Big) = -\infty. \tag{4.17}$$

It enables us by (4.15) to deduce that for all $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{M_n}{n} > \delta\right) = -\infty. \tag{4.18}$$

The same result is also true replacing M_n by $-M_n$ in (4.18) since M_n and $-M_n$ share the same distribution. Therefore, we find that

$$\frac{M_n}{n} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} 0. \tag{4.19}$$

A similar reasoning leads to the exponential convergence of N_n/n to 0, with speed b_n^2. Finally, we obtain (4.7) from (4.8) together with (4.10), (4.11), (4.14) and (4.19) which achieves the proof of Lemma 4.1. □

Corollary 4.2. *By virtue of Lemma 4.1 and under the same assumptions, we have the exponential convergence*

$$\frac{P_n}{n} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} \ell_1 \tag{4.20}$$

where $\ell_1 = \theta^* \ell$.

Proof. The proof of Corollary 4.2 is immediately derived from the following inequality,

$$\begin{aligned} \left|\frac{P_n}{n} - \theta^* \frac{S_n}{n}\right| &= \left|\frac{1}{1+\theta\rho}\frac{M_n}{n} + \frac{1}{1+\theta\rho}\frac{R_n(\theta)}{n} - \theta^*\frac{X_n^2}{n}\right|, \\ &\leq \frac{1}{1+\theta\rho}\frac{|M_n|}{n} + \frac{1}{1+\theta\rho}\frac{|R_n(\theta)|}{n} + |\theta^*|\frac{X_n^2}{n} \end{aligned} \tag{4.21}$$

with $R_n(\theta) = \theta\rho X_n X_{n-1} + \rho X_0(\varepsilon_0 - X_0)$. □

We are now in the position to prove Theorem 2.1. We shall make use of the following deviation principle for martingales established by Worms [21].

Theorem 4.3 (Worms). *Let (Y_n) be an adapted sequence with values in $\mathbb{R}^p$, and (V_n) a gaussian noise with variance $\sigma^2 > 0$. We suppose that (Y_n) satisfies, for some invertible square matrix C of order p and a speed sequence (b_n^2) such that $b_n^2 = o(n)$, the exponential convergence for any $\delta > 0$,*

$$\lim_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\left\|\frac{1}{n}\sum_{k=0}^{n-1} Y_k Y_k' - C\right\| > \delta\right) = -\infty. \tag{4.22}$$

Then, the sequence

$$\left(\frac{M_n}{b_n\sqrt{n}}\right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}^p$ of speed b_n^2 and good rate function

$$I(x) = \frac{1}{2\sigma^2} x' C^{-1} x \tag{4.23}$$

where (M_n) is the martingale given by

$$M_n = \sum_{k=1}^{n} Y_{k-1} V_k.$$

Proof. The proof of Theorem 4.3 is contained in the one of Theorem 5 of [21] with $d = 1$. □

Proof of Theorem 2.1. Let us consider the decomposition

$$\frac{\sqrt{n}}{b_n}\left(\widehat{\theta}_n-\theta^*\right)=\frac{\sqrt{n}}{b_n}\left(\frac{\sigma^2}{1+\theta\rho}\right)\frac{M_n}{\langle M\rangle_n}+\frac{\sqrt{n}}{b_n}\left(\frac{1}{1+\theta\rho}\right)\frac{R_n(\theta)}{S_{n-1}}, \tag{4.24}$$

that can be obtained by a straighforward calculation, where the remainder term $R_n(\theta)$ is defined in (4.21). First, by using the same methodology as in convergence (4.11), we obtain that for all $\delta>0$ and for a suitable $t>0$,

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{b_n^2}\log\mathbb{P}\left(\frac{X_0^2}{b_n\sqrt{n}}>\delta\right) &\leq \lim_{n\to\infty}\left(-t\delta\frac{\sqrt{n}}{b_n}\right)+\lim_{n\to\infty}\frac{1}{b_n^2}\log\mathbb{E}\Big[\exp(tX_0^2)\Big],\\ &= -\infty,\end{aligned} \tag{4.25}$$

since $b_n=o(\sqrt{n})$, and the same goes for any isolated term in (4.24) of order 2 whose numerator do not depend on n. Moreover, under the gaussian assumption on the driven noise (V_n), it is not hard to see that

$$\frac{1}{b_n\sqrt{n}}\max_{1\leq k\leq n}V_k^2 \overset{\text{superexp}}{\underset{b_n^2}{\longrightarrow}} 0. \tag{4.26}$$

As a matter of fact, for all $\delta>0$ and for all $t>0$,

$$\begin{aligned}\mathbb{P}\left(\max_{1\leq k\leq n}V_k^2\geq\delta b_n\sqrt{n}\right) &= \mathbb{P}\left(\bigcup_{k=1}^{n}\left\{V_k^2\geq\delta b_n\sqrt{n}\right\}\right) \leq \sum_{k=1}^{n}\mathbb{P}\Big(V_k^2\geq\delta b_n\sqrt{n}\Big),\\ &\leq n\exp\left(-t\delta b_n\sqrt{n}\right)\mathbb{E}\Big[\exp\left(tV_1^2\right)\Big].\end{aligned}$$

In addition, as soon as $0<t<1/(2\sigma^2)$, $\mathbb{E}\big[\exp(tV_1^2)\big]<\infty$. Consequently,

$$\begin{aligned}\frac{1}{b_n^2}\log\mathbb{P}\left(\max_{1\leq k\leq n}V_k^2\geq\delta b_n\sqrt{n}\right) &\leq \frac{\log n}{b_n^2}-\frac{t\delta\sqrt{n}}{b_n}+\frac{\log\mathbb{E}\Big[\exp\left(tV_1^2\right)\Big]}{b_n^2},\\ &\leq \frac{\sqrt{n}}{b_n}\left(\frac{\log n}{b_n\sqrt{n}}-t\delta+\frac{\log\mathbb{E}\Big[\exp\left(tV_1^2\right)\Big]}{b_n\sqrt{n}}\right)\end{aligned}$$

which clearly leads to (4.26). Furthermore, it follows from (1.1) that

$$\max_{1\leq k\leq n}X_k^2\leq\frac{1}{1-|\theta|}X_0^2+\left(\frac{1}{1-|\theta|}\right)^2\max_{1\leq k\leq n}\varepsilon_k^2, \tag{4.27}$$

as well as

$$\max_{1\leq k\leq n}\varepsilon_k^2\leq\frac{1}{1-|\rho|}\varepsilon_0^2+\left(\frac{1}{1-|\rho|}\right)^2\max_{1\leq k\leq n}V_k^2. \tag{4.28}$$

Then, we deduce from (4.25), (4.26), (4.27) and (4.28) that

$$\frac{1}{b_n\sqrt{n}}\max_{1\leq k\leq n}\varepsilon_k^2 \overset{\text{superexp}}{\underset{b_n^2}{\longrightarrow}} 0 \qquad\text{and}\qquad \frac{1}{b_n\sqrt{n}}\max_{1\leq k\leq n}X_k^2 \overset{\text{superexp}}{\underset{b_n^2}{\longrightarrow}} 0,$$

which of course imply the exponential convergence of $X_n^2/(b_n\sqrt{n})$ to 0, with speed b_n^2. Therefore, we obtain that

$$\frac{R_n(\theta)}{b_n\sqrt{n}} \overset{\text{superexp}}{\underset{b_n^2}{\longrightarrow}} 0. \tag{4.29}$$

We infer from Lemma 4.1 together with Lemma 4.1 of [22] that the following convergence is satisfied,

$$\frac{n}{S_n} \xrightarrow[b_n^2]{\text{superexp}} \frac{1}{\ell} \tag{4.30}$$

where $\ell > 0$ is given by (4.6). According to (4.29), the latter convergence and again Lemma 4.1 of [22], we deduce that

$$\frac{\sqrt{n}}{b_n} \left(\frac{1}{1+\theta\rho} \right) \frac{R_n(\theta)}{S_{n-1}} \xrightarrow[b_n^2]{\text{superexp}} 0. \tag{4.31}$$

Hence, we obtain from (4.30) that the same is true for

$$\frac{\sigma^2}{1+\theta\rho} \frac{M_n}{b_n\sqrt{n}} \left(\frac{n}{\langle M \rangle_n} - \frac{1}{\sigma^2 \ell} \right) \xrightarrow[b_n^2]{\text{superexp}} 0, \tag{4.32}$$

since Lemma 4.1 together with Theorem 4.3 with $p = 1$ directly show that $(M_n/(b_n\sqrt{n}))$ satisfies an LDP with speed b_n^2 and good rate function given, for all $x \in \mathbb{R}$, by

$$J(x) = \frac{x^2}{2\ell\sigma^2}. \tag{4.33}$$

As a consequence,

$$\frac{\sqrt{n}}{b_n} \left(\widehat{\theta}_n - \theta^* \right) \overset{\text{superexp}}{\underset{b_n^2}{\sim}} \frac{1}{\ell(1+\theta\rho)} \frac{M_n}{b_n\sqrt{n}}, \tag{4.34}$$

and this implies that both of them share the same LDP, see e.g. [6]. One shall now take advantage of the contraction principle [6] to establish that $(\sqrt{n}(\widehat{\theta}_n - \theta^*)/b_n)$ satisfies an LDP with speed b_n^2 and good rate function $I_\theta(x)$ given by (2.1). The contraction principle enables us to conclude that the good rate function of the LDP with speed b_n^2 associated with equivalence (4.34) is given by $I_\theta(x) = J(\ell(1+\theta\rho)x)$, that is

$$I_\theta(x) = \frac{x^2}{2\sigma_\theta^2},$$

which achieves the proof of Theorem 2.1. □

4.2. Proof of Theorem 2.2.

We need to introduce some more notations. For all $n \geq 2$, let

$$Q_n = \sum_{k=2}^{n} X_{k-2} V_k. \tag{4.35}$$

In addition, for all $n \geq 1$, denote

$$T_n = 1 + \theta^*\rho^* - \left(1 + \rho^*(\widehat{\theta}_n + \theta^*)\right) \frac{S_n}{S_{n-1}} + \left(2\rho^* + \widehat{\theta}_n + \theta^*\right) \frac{P_n}{S_{n-1}} - \frac{Q_n}{S_{n-1}}, \tag{4.36}$$

where S_n and P_n are respectively given by (4.3) and (4.5). Finally, for all $n \geq 0$, let

$$J_n = \sum_{k=0}^{n} \widehat{\varepsilon}_k^{\,2} \tag{4.37}$$

where the residual set $(\widehat{\varepsilon}_n)$ is given in (1.3). A set of additional technical tools has to be expounded to make the proof of Theorem 2.2 more tractable.

Corollary 4.4. *By virtue of Lemma 4.1 and under the same assumptions, we have the exponential convergence*

$$\frac{Q_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \ell_2$$

where $\ell_2 = ((\theta+\rho)\theta^* - \theta\rho)\ell$.

Proof. The proof of Corollary 4.4 immediately follows from the inequality,

$$\begin{aligned} \left| \frac{Q_n}{n} - ((\theta+\rho)\theta^* - \theta\rho)\frac{S_n}{n} \right| &= \left| \theta^* \frac{M_n}{n} + \frac{N_n}{n} + \frac{\xi_n^Q}{n} \right|, \\ &\leq |\theta^*|\frac{|M_n|}{n} + \frac{|N_n|}{n} + \frac{|\xi_n^Q|}{n} \end{aligned} \tag{4.38}$$

where ξ_n^Q is a residual made of isolated terms such that

$$\frac{\xi_n^Q}{n} \xrightarrow[b_n^2]{\text{superexp}} 0,$$

see e.g. the proof of Theorem 3.2 in [2] where more details are given on ξ_n^Q. □

Lemma 4.5. *Under the assumptions of Theorem 2.2, we have the exponential convergence*

$$A_n \xrightarrow[b_n^2]{\text{superexp}} A$$

where

$$A_n = \frac{n}{1+\theta\rho} \begin{pmatrix} \dfrac{1}{S_{n-1}} & 0 \\ \dfrac{T_n}{J_{n-1}} & -\dfrac{(\theta+\rho)}{J_{n-1}} \end{pmatrix}, \tag{4.39}$$

and

$$A = \frac{1}{\ell(1+\theta\rho)(1-(\theta^*)^2)} \begin{pmatrix} 1-(\theta^*)^2 & 0 \\ \theta\rho + (\theta^*)^2 & -(\theta+\rho) \end{pmatrix}. \tag{4.40}$$

Proof. Via (4.30), we directly obtain the exponential convergence,

$$\frac{1}{(1+\theta\rho)} \frac{n}{S_{n-1}} \xrightarrow[b_n^2]{\text{superexp}} \frac{1}{\ell(1+\theta\rho)}. \tag{4.41}$$

The combination of Lemma 4.1, Corollary 4.2, Corollary 4.4 and Lemma 4.1 of [22] shows, after a simple calculation, that

$$T_n \xrightarrow[b_n^2]{\text{superexp}} (\theta^*)^2 + \theta\rho. \tag{4.42}$$

Moreover, J_n given by (4.37) can be rewritten as

$$J_n = S_n - 2\widehat{\theta}_n P_n + \widehat{\theta}_n^{\,2} S_{n-1},$$

which leads, via Lemma 4.1 in [22], to

$$\frac{J_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \ell(1-(\theta^*)^2). \tag{4.43}$$

Convergences (4.42) and (4.43) imply

$$\left(\frac{n}{1+\theta\rho}\right) \frac{T_n}{J_{n-1}} \xrightarrow[b_n^2]{\text{superexp}} \frac{(\theta^*)^2 + \theta\rho}{\ell(1+\theta\rho)(1-(\theta^*)^2)}, \tag{4.44}$$

and finally,

$$\left(\frac{n}{1+\theta\rho}\right)\frac{\theta+\rho}{J_{n-1}} \xrightarrow[b_n^2]{\text{superexp}} \frac{\theta+\rho}{\ell(1+\theta\rho)(1-(\theta^*)^2)}. \tag{4.45}$$

Finally, (4.41) together with (4.44) and (4.45) achieve the proof of Lemma 4.5. □

Proof of Theorem 2.2. We shall make use of the decomposition

$$\frac{\sqrt{n}}{b_n}\begin{pmatrix}\widehat{\theta}_n-\theta^*\\ \widehat{\rho}_n-\rho^*\end{pmatrix} = \frac{1}{b_n\sqrt{n}}A_nZ_n + B_n, \tag{4.46}$$

where A_n is given by (4.39), $(Z_n)_{n\geq 0}$ is the 2-dimensional vector martingale given by

$$Z_n = \begin{pmatrix}M_n\\ N_n\end{pmatrix}, \tag{4.47}$$

and where the remainder term

$$B_n = \frac{1}{(1+\theta\rho)}\frac{\sqrt{n}}{b_n}\begin{pmatrix}\dfrac{R_n(\theta)}{S_{n-1}}\\ \dfrac{R_n(\rho)}{J_{n-1}}\end{pmatrix}. \tag{4.48}$$

The first component $R_n(\theta)$ is given in (4.21) while $R_n(\rho)$, whose definition may be found in the proof of Theorem 3.2 in [2], is made of isolated terms. Consequently, (4.25) and (4.29) are sufficient to ensure that

$$\frac{R_n(\theta)}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0 \qquad \text{and} \qquad \frac{R_n(\rho)}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0.$$

Therefore, we obtain that

$$B_n \xrightarrow[b_n^2]{\text{superexp}} 0. \tag{4.49}$$

In addition, it follows from Lemma 4.5 and Theorem 4.3 with $p=2$ that $(Z_n/(b_n\sqrt{n}))$ satisfies an LDP on $\mathbb{R}^2$ with speed b_n^2 and good rate function given, for all $x\in\mathbb{R}^2$, by

$$J(x) = \frac{1}{2\sigma^2}x'\Lambda^{-1}x, \tag{4.50}$$

where

$$\Lambda = \ell\begin{pmatrix}1 & \theta^*\\ \theta^* & 1\end{pmatrix}, \tag{4.51}$$

since we have the exponential convergence

$$\frac{\langle Z\rangle_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \sigma^2\Lambda \tag{4.52}$$

by application of Lemma 4.1 and Corollary 4.2. One observes that $\det(\Lambda) = \ell^2(1-(\theta^*)^2) > 0$ implying that Λ is invertible. As a consequence,

$$\frac{1}{b_n\sqrt{n}}(A_n-A)Z_n \xrightarrow[b_n^2]{\text{superexp}} 0, \tag{4.53}$$

and we deduce from (4.46) that

$$\frac{\sqrt{n}}{b_n}\begin{pmatrix}\widehat{\theta}_n-\theta^*\\ \widehat{\rho}_n-\rho^*\end{pmatrix} \underset{b_n^2}{\overset{\text{superexp}}{\sim}} \frac{1}{b_n\sqrt{n}}AZ_n. \tag{4.54}$$

This of course implies that both of them share the same LDP. The contraction principle [6] enables us to conclude that the rate function of the LDP on $\mathbb{R}^2$ with speed b_n^2 associated with equivalence (4.54) is given, for all $x \in \mathbb{R}^2$, by $K(x) = J(A^{-1}x)$, that is

$$K(x) = \frac{1}{2} x' \Gamma^{-1} x,$$

where $\Gamma = \sigma^2 A \Lambda A'$ is given by (1.10), and where we shall suppose that $\theta \neq -\rho$ to ensure that A is invertible. In particular, the latter result also implies that the good rate function of the LDP on $\mathbb{R}$ with speed b_n^2 associated with $(\sqrt{n}(\widehat{\rho}_n - \rho^*)/b_n)$ is given, for all $x \in \mathbb{R}$, by

$$I_\rho(x) = \frac{x^2}{2\sigma_\rho^2},$$

where σ_ρ^2 is the last element of the matrix Γ. This achieves the proof of Theorem 2.2. □

4.3. Proof of Theorem 2.3.

For all $n \geq 1$, denote by f_n the explosion coefficient associated with J_n given by (4.37), that is

$$f_n = \frac{J_n - J_{n-1}}{J_n} = \frac{\widehat{\varepsilon}_n^{\,2}}{J_n}. \tag{4.55}$$

It follows from decomposition (C.4) in [2] that

$$\frac{\sqrt{n}}{b_n}\Big(\widehat{D}_n - D^*\Big) = -2\frac{\sqrt{n}}{b_n}\Big(1 - f_n\Big)\Big(\widehat{\rho}_n - \rho^*\Big) + \frac{\sqrt{n}}{b_n}\zeta_n, \tag{4.56}$$

where the remainder term ζ_n is made of isolated terms. As before, we clearly have

$$\frac{\sqrt{n}}{b_n}\zeta_n \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} 0 \qquad \text{and} \qquad f_n \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} 0.$$

As a consequence,

$$\frac{\sqrt{n}}{b_n}\Big(\widehat{D}_n - D^*\Big) \underset{b_n^2}{\overset{\text{superexp}}{\sim}} -2\frac{\sqrt{n}}{b_n}\Big(\widehat{\rho}_n - \rho^*\Big), \tag{4.57}$$

and this implies that both of them share the same LDP. The contraction principle [6] enables us to conclude that the rate function of the LDP on $\mathbb{R}$ with speed b_n^2 associated with equivalence (4.57) is given, for all $x \in \mathbb{R}$, by $I_D(x) = I_\rho(-x/2)$, that is

$$I_D(x) = \frac{x^2}{2\sigma_D^2},$$

which achieves the proof of Theorem 2.3. □

4.4. Proofs of Theorem 3.1, Theorem 3.2 and Theorem 3.3.

We shall now propose a technical lemma ensuring that all results already proved under the gaussian assumption still hold under the Chen-Ledoux type condition.

Lemma 4.6. *Under* **(CL.1)**, **(CL.2)** *and* **(CL.3)**, *all exponential convergences of Lemma 4.1, Corollary 4.2, Corollary 4.4 and Lemma 4.5 still hold.*

Proof. Under **(CL.1)**, **(CL.2)** and **(CL.3)**, and following the same methodology as the one used to establish (4.29), we get

$$\frac{X_n^2}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0, \tag{4.58}$$

and Cauchy-Schwarz inequality implies that this is also the case for any isolated term of order 2, such as $X_n X_{n-1}/(b_n\sqrt{n})$. This allows us to control each remainder term. Note that **(CL.2)**, **(CL.3)** and (4.58) are obviously true for ε_0^4/n, X_0^4/n, ε_0^2/n, X_0^2/n and X_n^2/n, since $b_n\sqrt{n} = o(n)$. Moreover, if follows from Theorem 2.2 of [13] under **(CL.1)** with $a = 2$, that

$$\frac{L_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \sigma^2. \tag{4.59}$$

Furthermore, since (M_n) is a locally square integrable martingale, we infer from Theorem 2.1 of [3] that for all $x, y > 0$,

$$\mathbb{P}\Big(|M_n| > x, \langle M\rangle_n + [M]_n \leq y\Big) \leq 2\exp\left(-\frac{x^2}{2y}\right), \tag{4.60}$$

where the predictable quadratic variation $\langle M\rangle_n = \sigma^2 S_{n-1}$ is described in (4.3) and the total quadratic variation is given by $[M]_0 = 0$ and, for all $n \geq 1$, by

$$[M]_n = \sum_{k=1}^{n} X_{k-1}^2 V_k^2. \tag{4.61}$$

According to (4.60), we have for all $\delta > 0$ and a suitable $b > 0$,

$$\begin{aligned}
\mathbb{P}\left(\frac{|M_n|}{n} > \delta\right) &\leq \mathbb{P}\Big(|M_n| > \delta n, \langle M\rangle_n + [M]_n \leq nb\Big) + \mathbb{P}\Big(\langle M\rangle_n + [M]_n > nb\Big), \\
&\leq 2\exp\left(-\frac{n\delta^2}{2b}\right) + \mathbb{P}\Big(\langle M\rangle_n + [M]_n > nb\Big), \\
&\leq 2\max\left(\mathbb{P}\Big(\langle M\rangle_n + [M]_n > nb\Big), 2\exp\left(-\frac{n\delta^2}{2b}\right)\right).
\end{aligned}$$

Consequently,

$$\limsup_{n\to\infty} \frac{1}{b_n^2}\log\mathbb{P}\left(\frac{|M_n|}{n} > \delta\right) \leq \limsup_{n\to\infty} \frac{1}{b_n^2}\log\mathbb{P}\Big(\langle M\rangle_n + [M]_n > nb\Big). \tag{4.62}$$

We have for all $b > 0$,

$$\begin{aligned}
\mathbb{P}\Big(\langle M\rangle_n + [M]_n > nb\Big) &\leq \mathbb{P}\left(\langle M\rangle_n > \frac{nb}{2}\right) + \mathbb{P}\left([M]_n > \frac{nb}{2}\right), \\
&\leq 2\max\left(\mathbb{P}\left(\langle M\rangle_n > \frac{nb}{2}\right), \mathbb{P}\left([M]_n > \frac{nb}{2}\right)\right).
\end{aligned} \tag{4.63}$$

Moreover, for all $n \geq 1$, let us define

$$T_n = \sum_{k=0}^{n} X_k^4 \qquad \text{and} \qquad \Gamma_n = \sum_{k=1}^{n} V_k^4,$$

and note that we easily have the following inequality,

$$T_n \leq \alpha X_0^4 + \beta\varepsilon_0^4 + \beta\Gamma_n \quad \text{a.s.} \tag{4.64}$$

with $\alpha = 1 + (1 - |\theta|)^{-4}$ and $\beta = (1 - |\rho|)^{-4}(1 - |\theta|)^{-4}$. This implies that, for n large enough, one can find $\gamma > 0$ such that

$$T_n \leq \gamma \Gamma_n \quad \text{a.s.}$$

choosing for example $\gamma = 3\max(\alpha, \beta)$, under **(CL.2)** and **(CL.3)** for $a = 4$. According to Theorem 2.2 of [13] under **(CL.1)** with $a = 4$, we also have the exponential convergence,

$$\frac{\Gamma_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \tau^4, \tag{4.65}$$

where $\tau^4 = \mathbb{E}[V_1^4]$, leading, via Cauchy-Schwarz inequality and (4.64), to

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{[M]_n}{n} > \delta\right) &\leq \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{\Gamma_n}{n} > \frac{\delta}{\sqrt{\gamma}}\right), \\ &= -\infty, \end{aligned} \tag{4.66}$$

where $\delta > \tau^4\sqrt{\gamma}$. Exploiting (4.16) and (4.59), the same result can be achieved for $\langle M \rangle_n / n$ under **(CL.1)** with $a = 2$ and $\delta > \sigma^4 \gamma$. As a consequence, it follows from (4.63), (4.66) and the latter remark that

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{\langle M \rangle_n + [M]_n}{n} > b\right) = -\infty, \tag{4.67}$$

as soon as $b > \sigma^4\gamma + \tau^4\sqrt{\gamma}$. Therefore, the exponential convergence of M_n/n to 0 with speed b_n^2 is obtained via (4.62) and (4.67), that is, for all $\delta > 0$ and $b > \sigma^4\gamma + \tau^4\sqrt{\gamma}$,

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{|M_n|}{n} > \delta\right) = -\infty. \tag{4.68}$$

The same obviously holds for N_n/n. Following the same lines as in the proofs of Lemma 4.1, Corollary 4.2, Corollary 4.4 and Lemma 4.5, hypothesis **(CL.2)** and **(CL.3)** with $a = 4$ together with exponential convergences (4.58), (4.59) and (4.68) are sufficient to achieve the proof of Lemma 4.6. □

Let us introduce a simplified version of Puhalskii's result [19] applied to a sequence of martingale differences, and two technical lemmas that shall help us to prove our results.

Theorem 4.7 (Puhalskii). *Let $(m_j^n)_{1\leq j\leq n}$ be a triangular array of martingale differences with values in $\mathbb{R}^d$, with respect to the filtration $(\mathcal{F}_n)_{n\geq 1}$. Let (b_n) be a sequence of real numbers satisfying* (1.13). *Suppose that there exists a symmetric positive-semidefinite matrix Q such that*

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\Big[m_k^n (m_k^n)' \Big| \mathcal{F}_{k-1}\Big] \xrightarrow[b_n^2]{\text{superexp}} Q. \tag{4.69}$$

Suppose that there exists a constant $c > 0$ such that, for each $1 \leq k \leq n$,

$$|m_k^n| \leq c\frac{\sqrt{n}}{b_n} \quad \text{a.s.} \tag{4.70}$$

Suppose also that, for all $a > 0$, we have the exponential Lindeberg's condition

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\Big[|m_k^n|^2 \mathrm{I}_{\left\{|m_k^n| \geq a\frac{\sqrt{n}}{b_n}\right\}} \Big| \mathcal{F}_{k-1}\Big] \xrightarrow[b_n^2]{\text{superexp}} 0. \tag{4.71}$$

Then, the sequence

$$\left(\frac{1}{b_n\sqrt{n}}\sum_{k=1}^{n} m_k^n\right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}^d$ *with speed* b_n^2 *and good rate function*

$$\Lambda^*(v) = \sup_{\lambda\in\mathbb{R}^d}\left(\lambda' v - \frac{1}{2}\lambda' Q\lambda\right).$$

In particular, if Q *is invertible,*

$$\Lambda^*(v) = \frac{1}{2}v' Q^{-1} v. \tag{4.72}$$

Proof. The proof of Theorem 4.7 is contained e.g. in the proof of Theorem 3.1 in [19]. □

Lemma 4.8. *Under* **(CL.1)**, **(CL.2)** *and* **(CL.3)** *with* $a = 2$, *we have for all* $\delta > 0$,

$$\limsup_{R\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k^2 \mathrm{I}_{\{|X_k|>R\}} > \delta\right) < 0.$$

Remark 4.1. *Lemma 4.8 implies that the exponential Lindeberg's condition given by* (4.71) *is satisfied.*

Proof. We introduce the empirical measure associated with the geometric ergodic Markov chain $(X_n)_{n\geq 0}$,

$$\Lambda_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{X_k}, \tag{4.73}$$

with invariant probability measure denoted by μ. It is well-known that the sequence (Λ_n) satisfies the upper bound of the moderate deviations, see e.g. [8] for more details. Let us define, for $f(x) = x^2$, the following truncations,

$$f^{(R)}(x) = f(x)\min\Big(1, \big(f(x) - (R-1)\big)_+\Big) \quad \text{and} \quad \widetilde{f}^{(R)}(x) = \min\Big(f^{(R)}(x), R\Big).$$

Thus, we have

$$0 \leq f(x)\mathrm{I}_{\{f(x)\geq R\}} \leq f^{(R)}(x) \leq f(x),$$

and, as a consequence,

$$0 \leq \Lambda_n\Big(f\mathrm{I}_{\{f\geq R\}}\Big) \leq \Lambda_n\Big(f^{(R)} - \widetilde{f}^{(R)}\Big) + \Lambda_n\Big(\widetilde{f}^{(R)}\Big) - \mu\Big(\widetilde{f}^{(R)}\Big) + \mu\Big(\widetilde{f}^{(R)}\Big).$$

We also have

$$f^{(R)} - \widetilde{f}^{(R)} = \Big(f^{(R)} - R\Big)\mathrm{I}_{\{f^{(R)}\geq R\}} \leq \Big(f - R\Big)\mathrm{I}_{\{f\geq R\}} = f - \Big(f\wedge R\Big).$$

For $\delta > 0$, the functions $\widetilde{f}^{(R)}$ and $f - (f\wedge R)$ are continuous and bounded by f which is μ-integrable, and they converge to 0 as R goes to infinity. By Lebesgue's Theorem, there exists $R > 0$ large enough such that $\mu(\widetilde{f}^{(R)}) + \mu(f - (f\wedge R)) < \delta/4$. Thus,

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k^2\mathrm{I}_{\{X_k^2\geq R\}} > \delta\right) \leq\ & \mathbb{P}\Big(\Lambda_n(f) - \mu(f) > \delta/4\Big) + \mathbb{P}\Big(\Lambda_n(f\wedge R) - \mu(f\wedge R) > \delta/4\Big)\\ &+ \mathbb{P}\Big(\Lambda_n(\widetilde{f}^{(R)}) - \mu(\widetilde{f}^{(R)}) > \delta/4\Big).\end{aligned} \tag{4.74}$$

From Lemma 4.6, we have that for all $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\Lambda_n(f) - \mu(f) > \delta\Big) = -\infty.$$

By the upper bound of the moderate deviation principle for the sequence (Λ_n) given in [8], we obtain that

$$\limsup_{R\to\infty} \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\Lambda_n(f\wedge R) - \mu(f\wedge R) > \delta\Big) = -\infty,$$

and

$$\limsup_{R\to\infty} \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\Lambda_n(\widetilde{f}^{(R)}) - \mu(\widetilde{f}^{(R)}) > \delta\Big) = -\infty,$$

which, via inequality (4.74), achieves the proof of Lemma 4.8. Note that Remark 4.1 is immediately derived from the latter proof, see e.g. [22] for more details. □

Lemma 4.9. *Under* **(CL.1)**, **(CL.2)** *and* **(CL.3)**, *the sequence*

$$\left(\frac{M_n}{b_n\sqrt{n}}\right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}$ *with speed* b_n^2 *and good rate function*

$$J(x) = \frac{x^2}{2\ell\sigma^2} \tag{4.75}$$

where ℓ *is given by* (4.6).

Proof. From now on, in order to apply Puhalskii's result for the moderate deviations for martingales, we introduce the following modification of the martingale $(M_n)_{n\geq 0}$, for $r > 0$ and $R > 0$,

$$M_n^{(r,R)} = \sum_{k=1}^{n} X_{k-1}^{(r)} V_k^{(R)} \tag{4.76}$$

where, for all $1 \leq k \leq n$,

$$X_k^{(r)} = X_k \mathrm{I}_{\left\{|X_k|\leq r\frac{\sqrt{n}}{b_n}\right\}} \quad \text{and} \quad V_k^{(R)} = V_k \mathrm{I}_{\left\{|V_k|\leq R\right\}} - \mathbb{E}\Big[V_k \mathrm{I}_{\left\{|V_k|\leq R\right\}}\Big]. \tag{4.77}$$

Then, we have to prove that for all $r > 0$ the sequence $(M_n^{(r,R)})$ is an exponentially good approximation of (M_n) as R goes to infinity, see e.g. Definition 4.2.14 in [6]. This approximation, in the sense of the large deviations, is described by the following convergence, for all $r > 0$ and all $\delta > 0$,

$$\limsup_{R\to\infty} \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{|M_n - M_n^{(r,R)}|}{b_n\sqrt{n}} > \delta\right) = -\infty. \tag{4.78}$$

From Lemma 4.6, and since $\langle M\rangle_n = \sigma^2 S_{n-1}$, we have

$$\frac{\langle M\rangle_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \sigma^2\ell. \tag{4.79}$$

From Lemma 4.6 and Remark 4.1, we also have for all $r > 0$,

$$\frac{1}{n}\sum_{k=0}^{n} X_k^2 \mathrm{I}_{\left\{|X_k|>r\frac{\sqrt{n}}{b_n}\right\}} \xrightarrow[b_n^2]{\text{superexp}} 0. \tag{4.80}$$

We introduce the following notations,

$$\sigma_R^2 = \mathbb{E}\left[(V_1^{(R)})^2\right] \qquad \text{and} \qquad S_n^{(r)} = \sum_{k=0}^{n} (X_k^{(r)})^2.$$

Then, we easily transfer properties (4.79) and (4.80) to the truncated martingale $(M_n^{(r,R)})_{n\geq 0}$. We have for all $R > 0$ and all $r > 0$,

$$\frac{\langle M^{(r,R)}\rangle_n}{n} = \sigma_R^2 \frac{S_{n-1}^{(r)}}{n} = -\sigma_R^2 \left(\frac{S_{n-1}}{n} - \frac{S_{n-1}^{(r)}}{n}\right) + \sigma_R^2 \frac{S_{n-1}}{n} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} \sigma_R^2 \ell$$

which ensures that (4.69) is satisfied for the martingale $(M_n^{(r,R)})_{n\geq 0}$. Note also that Lemma 4.6 and Remark 4.1 work for the martinagle $(M_n^{(r,R)})_{n\geq 0}$. So, for all $r > 0$, the exponential Lindeberg's condition and thus (4.71) are satisfied for $(M_n^{(r,R)})_{n\geq 0}$. By Theorem 4.7, we deduce that $(M_n^{(r,R)}/b_n\sqrt{n})$ satisfies an LDP on $\mathbb{R}$ with speed b_n^2 and good rate function

$$J_R(x) = \frac{x^2}{2\sigma_R^2 \ell}. \tag{4.81}$$

It will be possible to drive the moderate deviations result for the martingale $(M_n)_{n\geq 0}$ by proving relation (4.78). For that matter, let us now introduce the following decomposition,

$$M_n - M_n^{(r,R)} = L_n^{(r)} + F_n^{(r,R)}$$

where

$$L_n^{(r)} = \sum_{k=1}^{n} \left(X_{k-1} - X_{k-1}^{(r)}\right) V_k \quad \text{and} \quad F_n^{(r,R)} = \sum_{k=1}^{n} \left(V_k - V_k^{(R)}\right) X_{k-1}^{(r)}.$$

One has to show that for all $r > 0$,

$$\frac{L_n^{(r)}}{b_n\sqrt{n}} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} 0, \tag{4.82}$$

and, for all $r > 0$ and all $\delta > 0$, that

$$\limsup_{R\to\infty} \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{|F_n^{(r,R)}|}{b_n\sqrt{n}} > \delta\right) = -\infty. \tag{4.83}$$

On the one hand, note that for any $\eta > 0$,

$$\sum_{k=0}^{n} |X_k|^{2+\eta} \leq \alpha |X_0|^{2+\eta} + \beta |\varepsilon_0|^{2+\eta} + \beta \sum_{k=1}^{n} |V_k|^{2+\eta} \quad \text{a.s.}$$

with $\alpha = 1 + (1 - |\theta|)^{-(2+\eta)}$ and $\beta = (1 - |\rho|)^{-(2+\eta)}(1 - |\theta|)^{-(2+\eta)}$. This implies that, for n large enough, one can find $\gamma > 0$ such that

$$\sum_{k=0}^{n} |X_k|^{2+\eta} \leq \gamma \sum_{k=1}^{n} |V_k|^{2+\eta} \quad \text{a.s.} \tag{4.84}$$

taking for example $\gamma = 3\max(\alpha, \beta)$, under **(CL.2)** and **(CL.3)** for $a = 2 + \eta$. Thus,

$$\begin{aligned} \frac{|L_n^{(r)}|}{b_n\sqrt{n}} &= \frac{1}{b_n\sqrt{n}} \left| \sum_{k=1}^n X_{k-1} \mathrm{I}_{\left\{|X_{k-1}|>r\frac{\sqrt{n}}{b_n}\right\}} V_k \right|, \\ &\leq \frac{1}{b_n\sqrt{n}} \left(r\frac{\sqrt{n}}{b_n} \right)^{-\eta} \left(\sum_{k=1}^n |X_{k-1}|^{2+\eta} \right)^{1/2} \left(\sum_{k=1}^n V_k^2 |X_{k-1}|^{\eta} \right)^{1/2}, \\ &\leq \lambda(r,\eta,\gamma) \left(\frac{b_n}{\sqrt{n}} \right)^{\eta-1} \frac{1}{n} \sum_{k=1}^n |V_k|^{2+\eta} \quad \text{a.s.} \end{aligned} \tag{4.85}$$

by virtue of (4.84) and Hölder's inequality, where $\lambda(r, \eta, \gamma) > 0$ can be evaluated under suitable assumptions of moment on (V_n). As a consequence, for all $\delta > 0$,

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{|L_n^{(r)}|}{b_n\sqrt{n}} > \delta \right) &\leq \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^n |V_k|^{2+\eta} > \frac{\delta}{\lambda(r,\eta,\gamma)} \left(\frac{\sqrt{n}}{b_n} \right)^{\eta-1} \right), \\ &= -\infty, \end{aligned} \tag{4.86}$$

as soon as $\eta > 1$, under **(CL.1)** with $a = 2 + \eta$. We deduce that

$$\frac{L_n^{(r)}}{b_n\sqrt{n}} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} 0, \tag{4.87}$$

which achieves the proof of (4.82), under **(CL.1)**, **(CL.2)** and **(CL.3)** for $a > 3$. On the other hand, $(F_n^{(r,R)})_{n\geq 0}$ is a locally square-integrable real martingale whose predictable quadratic variation is given by $\langle F^{(r,R)} \rangle_0 = 0$ and, for all $n \geq 1$, by

$$\langle F^{(r,R)} \rangle_n = \mathbb{E}\left[\left(V_1 - V_1^{(R)} \right)^2 \right] S_{n-1}^{(r)}.$$

To prove (4.83), we will use Theorem 1 of [7]. For R large enough and all $k \geq 1$, we have

$$\begin{aligned} \mathbb{P}\left(\left| X_{k-1}^{(r)} \left(V_k - V_k^{(R)} \right) \right| > b_n\sqrt{n} \,\Big|\, \mathcal{F}_{k-1} \right) &\leq \mathbb{P}\left(\left| V_k - V_k^{(R)} \right| > \frac{b_n^2}{r} \right), \\ &= \mathbb{P}\left(\left| V_1 - V_1^{(R)} \right| > \frac{b_n^2}{r} \right) = 0. \end{aligned}$$

This implies that

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log \left(n \operatorname*{ess\,sup}_{k\geq 1} \mathbb{P}\left(\left| X_{k-1}^{(r)} \left(V_k - V_k^{(R)} \right) \right| > b_n\sqrt{n} \,\Big|\, \mathcal{F}_{k-1} \right) \right) = -\infty. \tag{4.88}$$

For all $\gamma > 0$ and all $\delta > 0$, we obtain from Lemma 4.8 and Remark 4.1, that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^n \left(X_{k-1}^{(r)} \right)^2 \mathrm{I}_{\left\{|X_{k-1}^{(r)}|>\gamma\frac{\sqrt{n}}{b_n}\right\}} > \delta \right) \leq \\ &\qquad \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^n X_{k-1}^2 \mathrm{I}_{\left\{|X_{k-1}|>\gamma\frac{\sqrt{n}}{b_n}\right\}} > \delta \right) = -\infty. \end{aligned}$$

Finally, from Lemma 4.6, Lemma 4.8 and Remark 4.1, it follows that

$$\frac{\langle F^{(r,R)}\rangle_n}{n} = Q_R \frac{S_{n-1}^{(r)}}{n} = -Q_R \left(\frac{S_{n-1}}{n} - \frac{S_{n-1}^{(r)}}{n} \right) + Q_R \frac{S_{n-1}}{n} \underset{b_n^2}{\overset{\text{superexp}}{\longrightarrow}} Q_R \ell$$

where

$$Q_R = \mathbb{E}\left[\left(V_1 - V_1^{(R)} \right)^2 \right],$$

and ℓ is given by (4.6). Moreover, it is clear that Q_R converges to 0 as R goes to infinity. In light of foregoing, we infer from Theorem 1 of [7] that $(F_n^{(r,R)}/(b_n\sqrt{n}))$ satisfies an LDP on $\mathbb{R}$ of speed b_n^2 and good rate function

$$I_R(x) = \frac{x^2}{2Q_R\ell}.$$

In particular, this implies that for all $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{|F_n^{(r,R)}|}{b_n\sqrt{n}} > \delta \right) \leq -\frac{\delta^2}{2Q_R\ell}, \tag{4.89}$$

and letting R go to infinity clearly leads to the end of the proof of (4.83). We are able to conclude now that $(M_n^{(r,R)}/(b_n\sqrt{n}))$ is an exponentially good approximation of $(M_n/(b_n\sqrt{n}))$. By application of Theorem 4.2.16 in [6], we find that $(M_n/(b_n\sqrt{n}))$ satisfies an LDP on $\mathbb{R}$ with speed b_n^2 and good rate function

$$\widetilde{J}(x) = \sup_{\delta>0} \liminf_{R\to\infty} \inf_{z\in B_{x,\delta}} J_R(z),$$

where J_R is given in (4.81) and $B_{x,\delta}$ denotes the ball $\{z : |z-x| < \delta\}$. The identification of the rate function $\widetilde{J} = J$, where J is given in (4.75) is done easily, which concludes the proof of Lemma 4.9. □

Remark 4.2. *If we suppose that* **(CL.1)** *holds with $a > 2$, then the exponential Lindeberg's condition in Lemma 4.8 is easier to establish. Indeed, using (4.84), it follows that*

$$\left(r\frac{\sqrt{n}}{b_n} \right)^\eta \sum_{k=1}^n X_{k-1}^2 \mathrm{I}_{\left\{|X_{k-1}|>r\frac{\sqrt{n}}{b_n}\right\}} \leq \sum_{k=1}^n |X_{k-1}|^{2+\eta} \leq \gamma \sum_{k=1}^n |V_k|^{2+\eta},$$

for n large enough and $\eta > 0$, leading to

$$\mathbb{P}\left(\frac{1}{n} \sum_{k=1}^n X_{k-1}^2 \mathrm{I}_{\left\{|X_{k-1}|>r\frac{\sqrt{n}}{b_n}\right\}} > \delta \right) \leq \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^n |V_k|^{2+\eta} > \frac{\delta}{\gamma} \left(r\frac{\sqrt{n}}{b_n} \right)^\eta \right).$$

Lemma 4.10. *Under* **(CL.1)**, **(CL.2)** *and* **(CL.3)**, *the sequence*

$$\left(\frac{1}{b_n\sqrt{n}} \begin{pmatrix} M_n \\ N_n \end{pmatrix} \right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}^2$ with speed b_n^2 and good rate function

$$J(x) = \frac{1}{2\sigma^2} x' \Lambda^{-1} x \tag{4.90}$$

where Λ is given by (4.51).

Proof. We follow the same approach as in the proof of Lemma 4.9. We shall consider the 2-dimensional vector martingale $(Z_n)_{n\geq 0}$ defined in (4.47). In order to apply Theorem 4.7, we introduce the following truncation of the martingale $(Z_n)_{n\geq 0}$, for $r > 0$ and $R > 0$,

$$Z_n^{(r,R)} = \begin{pmatrix} M_n^{(r,R)} \\ N_n^{(r,R)} \end{pmatrix}$$

where $M_n^{(r,R)}$ is given in (4.76) and where $N_n^{(r,R)}$ is defined in the same manner, that is, for all $n \geq 2$,

$$N_n^{(r,R)} = \sum_{k=2}^{n} X_{k-2}^{(r)} V_k^{(R)} \tag{4.91}$$

with $X_n^{(r)}$ and $V_n^{(R)}$ given by (4.77). The exponential convergence (4.52) still holds, by virtue of Lemma 4.6, which immediately implies hypothesis (4.69). On top of that, Lemma 4.8 ensures that, for all $r > 0$,

$$\frac{1}{n}\sum_{k=0}^{n} X_k^2 \mathrm{I}_{\left\{|X_k|>r\frac{\sqrt{n}}{b_n}\right\}} \xrightarrow[b_n^2]{\text{superexp}} 0, \tag{4.92}$$

justifying hypothesis (4.71). Via Theorem 4.7, $(Z_n^{(r,R)}/(b_n\sqrt{n}))$ satisfies an LDP on $\mathbb{R}^2$ with speed b_n^2 and good rate function J_R given by

$$J_R(x) = \frac{1}{2\sigma_R^2} x' \Lambda^{-1} x. \tag{4.93}$$

Finally, it is straightforward to prove that $(Z_n^{(r,R)}/(b_n\sqrt{n}))$ is an exponentially good approximation of $(Z_n/(b_n\sqrt{n}))$. By application of Theorem 4.2.16 in [6], we deduce that $(Z_n/(b_n\sqrt{n}))$ satisfies an LDP on $\mathbb{R}^2$ with speed b_n^2 and good rate function given by

$$\widetilde{J}(x) = \sup_{\delta>0} \liminf_{R\to\infty} \inf_{z\in B_{x,\delta}} J_R(z),$$

where J_R is given in (4.93) and $B_{x,\delta}$ denotes the ball $\{z : |z-x| < \delta\}$. The identification of the rate function $\widetilde{J} = J$ is done easily, which concludes the proof of Lemma 4.10. $\square$

Proofs of Theorem 3.1, Theorem 3.2 and Theorem 3.3. The residuals appearing in the decompositions (4.24), (4.46) and (4.56) still converge exponentially to zero under **(CL.1)**, **(CL.2)** and **(CL.3)**, with speed b_n^2, as it was already proved. Therefore, for a better readability, we may skip the most accessible parts of these proofs whose development merely consists in following the same lines as those in the proofs of Theorem 2.1, Theorem 2.2 and Theorem 2.3, taking advantage of Lemma 4.9 and Lemma 4.10, and applying the contraction principle given e.g. in [6]. $\square$

Acknowledgments. *The authors thank Bernard Bercu and Arnaud Guillin for all their advices and suggestions during the preparation of this work.*

References

[1] Arcones, M. A. The large deviation principle of empirical processes. *Preprint* (2001).
[2] Bercu, B., and Proïa, F. A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process. *arXiv 1104.3328v1. In revision. ESAIM Probab. Stat.* (2011).

[3] Bercu, B., and Touati, A. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab. 18, no.5* (2008), 1848–1869.

[4] Chen, X. Moderate deviations for m-dependent random variables with Banach space value. *Statis. and Probab. Letters. 35* (1998), 123–134.

[5] Dembo, A. Moderate deviations for martingales with bounded jumps. *Electron. Comm. Probab. 1, no. 3* (1996), 11–17.

[6] Dembo, A., and Zeitouni, O. *Large deviations techniques and applications, second edition*, vol. 38 of *Applications of Mathematics*. Springer, 1998.

[7] Djellout, H. Moderate deviations for martingale differences and applications to ϕ-mixing sequences. *Stoch. Stoch. Rep. 73, 1-2* (2002), 37–63.

[8] Djellout, H., and Guillin, A. Moderate deviations for Markov chains with atom. *Stochastic Process. Appl. 95, no. 2* (2001), 203–217.

[9] Durbin, J. Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. *Econometrica 38* (1970), 410–421.

[10] Durbin, J., and Watson, G. S. Testing for serial correlation in least squares regression. I. *Biometrika 37* (1950), 409–428.

[11] Durbin, J., and Watson, G. S. Testing for serial correlation in least squares regression. II. *Biometrika 38* (1951), 159–178.

[12] Durbin, J., and Watson, G. S. Testing for serial correlation in least squares regession. III. *Biometrika 58* (1971), 1–19.

[13] Eichelsbacher, P., and Löwe, M. Moderate deviations for i.i.d. random variables. *ESAIM Probab. Stat. 7* (2003), 209–218.

[14] Inder, B. A. An approximation to the null distribution of the Durbin-Watson statistic in models containing lagged dependent variables. *Econometric Theory 2* (1986), 413–428.

[15] King, M. L., and Wu, P. X. Small-disturbance asymptotics and the Durbin-Watson and related tests in the dynamic regression model. *J. Econometrics 47* (1991), 145–152.

[16] Ledoux, M. Sur les déviations modérées des sommes de variables aléatoires vectorielles indépendantes de même loi. *Ann. Inst. Henri-Poincaré. 35* (1992), 123–134.

[17] Malinvaud, E. Estimation et prévision dans les modèles économiques autorégressifs. *Review of the International Institute of Statistics 29* (1961), 1–32.

[18] Nerlove, M., and Wallis, K. F. Use of the Durbin-Watson statistic in inappropriate situations. *Econometrica 34* (1966), 235–238.

[19] Puhalskii, A. Large deviations of semimartingales: a maxingale problem approach. I. Limits as solutions to a maxingale problem. *Stoch. Stoch. Rep. 61* (1997, no. 3-4), 141–243.

[20] Stocker, T. On the asymptotic bias of OLS in dynamic regression models with autocorrelated errors. *Statist. Papers 48* (2007), 81–93.

[21] Worms, J. Moderate deviations for stable Markov chains and regression models. *Electron. J. Probab. 4, no. 8* (1999), 1–28.

[22] Worms, J. Principes de déviations modérées pour des martingales et applications statistiques. *Thèse de Doctorat à l'Université Marne-la-Vallée.* (2000).

[23] Worms, J. Moderate deviations of some dependent variables. I. Martingales. *Math. Methods Statist. 10, no. 1* (2001), 38–72.

[24] Worms, J. Moderate deviations of some dependent variables. II. Some kernel estimators. *Math. Methods Statist. 10, no. 2* (2001), 161–193.

E-mail address: `Valere.Bitsekipenda@math.univ-bpclermont.fr`
E-mail address: `Hacene.Djellout@math.univ-bpclermont.fr`

Laboratoire de Mathématiques, CNRS UMR 6620, Université Blaise Pascal, Avenue des Landais, 63177 Aubière, France.

E-mail address: `Frederic.Proia@inria.fr`

Université Bordeaux 1, Institut de Mathématiques de Bordeaux, UMR 5251, and INRIA Bordeaux, team ALEA, 351 Cours de la Libération, 33405 Talence cedex, France.